Synthetic promoter design in Escherichia coli based on a deep generative network

Year: 2020 | *Citations: 153* | *Authors: Ye Wang, Haochen Wang, Lei Wei*

Abstract

Abstract Promoter design remains one of the most important considerations in metabolic engineering and synthetic biology applications. Theoretically, there are 450 possible sequences for a 50-nt promoter, of which naturally occurring promoters make up only a small subset. To explore the vast number of potential sequences, we report a novel AI-based framework for de novo promoter design in Escherichia coli. The model, which was guided by sequence features learned from natural promoters, could capture interactions between nucleotides at different positions and design novel synthetic promoters in silico. We combined a deep generative model that guides the search for artificial sequences with a predictive model to preselect the most promising promoters. The AI-designed promoters were optimized based on the promoter activity in E. coli and the predictive model. After two rounds of optimization, up to 70.8% of the AI-designed promoters were experimentally demonstrated to be functional, and few of them shared significant sequence similarity with the E. coli genome. Our work provided an end-to-end approach to the de novo design of novel promoter elements, indicating the potential to apply deep learning methods to de novo genetic element design.